A computational model of generalization in distributional learning: the role of phonetic variability across segment classes

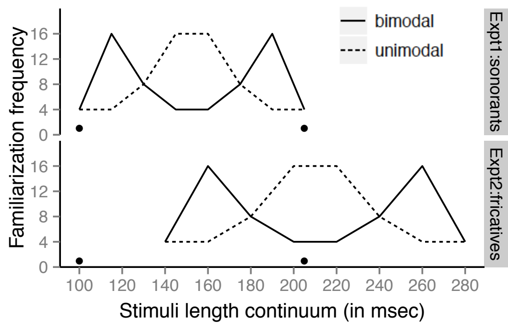
Bozena Pajak,[a] Klinton Bicknell,[b] and Roger Levy[b]
[a] University of Rochester, [b] UC San Diego

One of the central problems in language acquisition is how phonetic categories are learned, involving mapping phonetic tokens that vary along continuous dimensions onto discrete categories. This task may be facilitated by languages' extensive re-use of a set of phonetic dimensions (Clements 2003), because learning one distinction (e.g., /b/-/p/ varying along the voice onset time (VOT) dimension) might help learn analogous distinctions (e.g., /d/-/k/, /g/-/k/). However, the difficulty here is variability in how exactly these distinctions are implemented within classes: VOT values are consistently lower for labials than for velars (Lisker & Abramson 1970); and the durations of singleton and geminate consonants are shorter for nasals than for voiceless fricatives (Giovanardi & Di Benedetto 1998, Mattei & Di Benedetto 2000). Given this variability, how do people use information about one distinction when learning an analogous distinction for another class? Here we propose that people's linguistic knowledge includes the expectation that category types in any language (such as voiced and voiceless, or singleton and geminate) can be shared across classes, but that different classes can implement them in idiosyncratic ways. We formalize this account in a hierarchical Bayesian model, and present simulations that reproduce key features of human performance in behavioral experiments.

Our prior work (Pajak & Levy 2011) tested this proposal by giving adult native English speakers distributional information about one segment class along the length continuum (see 1), and probing their expectations about the number of categories in that class and another untrained class (by asking for judgments about tokens at the endpoints of the continua, as illustrated in 2). The results showed that learners infer the number of categories from frequency distributions: they infer two categories when the distribution is bimodal, and a single category when the distribution is unimodal (3, 'trained' condition). In addition, learners use the information about the trained class to make inferences about the untrained class (3, 'untrained' condition): they are more likely to accept length-based distinctions for fricatives after learning the distinction for sonorants (Expt. 1), and vice versa (Expt. 2). Crucially, this generalization occurs both (a) when each class implements the distinction in exactly the same way (with the same absolute durations; Expt. 1), and (b) when the classes differ in how the shared distinction type is implemented (the absolute durations of the untrained class are shifted relative to the trained class; Expt. 2).

Our model reproduces this behavior by using distributional information to learn phonetic categories via nonparametric Bayesian inference, where categories are modeled as Gaussian distributions (cf. Feldman et al. 2009). We use a Hierarchical Dirichlet Process prior to allow for the learning of an unbounded number of categories, which may be shared across segment classes. For example, the model can learn a 'geminate' category that is shared across sonorant and fricative sound classes. In order to assess the role of variability, we compare two versions of the model: (a) a basic model, which does not explicitly allow variability across segment classes, and (b) an extended model that allows for one type of idiosyncratic implementation of categories across classes, learnable class-specific 'offsets' by which categories in a class are shifted. Our results demonstrate that the basic model cannot reproduce the human pattern (see 4): it learns from distributional cues and generalizes the learned categories to an untrained class when the absolute durations of both classes are aligned (as in Expt. 1), but it fails to generalize when the untrained class has the same category structure but different absolute values (as in Expt. 2). The expanded model, on the other hand, reproduces the full generalization pattern (see 5), suggesting that allowing for variability across segment classes may be necessary to account for human learning. Taken together, this work suggests that language learners have implicit knowledge of the ways that sound classes can vary, and that they can harness this knowledge to take advantage of the underlying similarities between sound classes with differing surface representations.

(1) Experiment 1 & 2 training (Pajak & Levy 2011)
The four points show the values of the untrained datapoints.



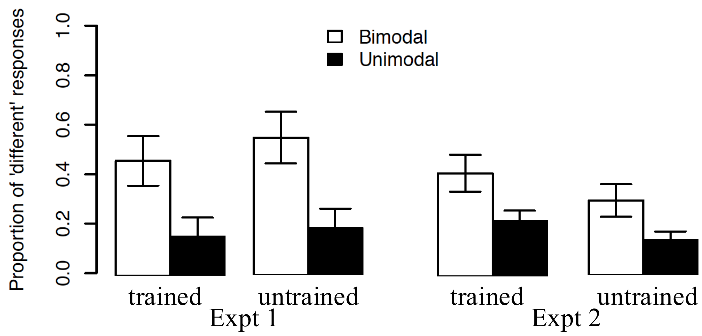
(2) Experiment 1 & 2 test (Pajak & Levy 2011)

“Are these two different words in this language or two repetitions of the same word?”

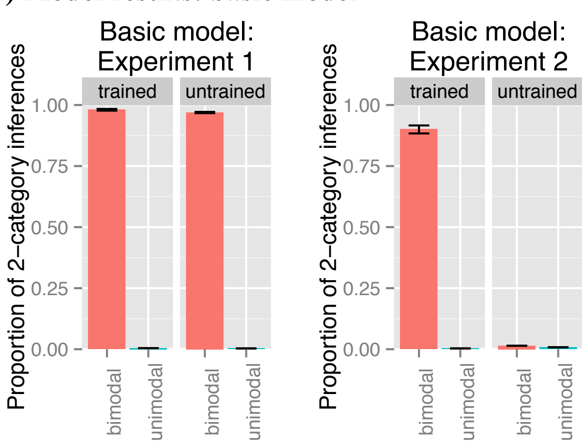
“different” trials
Examples: [ama]-[amma], [asa]-[assa]

	Expt 1	Expt 2
trained	(sonorants) 100ms -205ms	(fricatives) 140ms-280ms
untrained	(fricatives) 100ms-205ms	(sonorants) 100ms-205ms

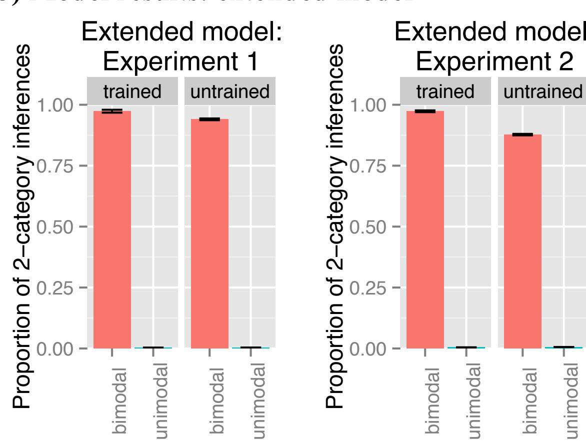
(3) Experiment 1 & 2 results (Pajak & Levy 2011)



(4) Model results: basic model



(5) Model results: extended model



References

Clements, G. N. (2003). Feature economy in sound systems. *Phonology*, *20*, 287–333.

Feldman, N. H., Griffiths, T. L., & Morgan, J. L. (2009). Learning phonetic categories by learning a lexicon. In *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 2208–2213). Austin, TX: Cognitive Science Society.

Giovanardi, M., & Di Benedetto, M.-G. (1998). Acoustic analysis of singleton and geminate fricatives in Italian. *The European Journal of Language and Speech (EACL/ESCA/ELSNET)*, *1998*, 1-13.

Lisker, L., & Abramson, A. S. (1970). The voicing dimensions: Some experiments in comparative phonetics. In *Proceedings of the Sixth International Congress of Phonetic Sciences*. Prague: Academia.

Mattei, M., & Di Benedetto, M.-G. (2000). Acoustic analysis of singleton and geminate nasals in Italian. *The European Journal of Language and Speech (EACL/ESCA/ELSNET)*, *2000*, 1–11.

Pajak, B., & Levy, R. (2011). How abstract are phonological representations? Evidence from distributional perceptual learning. In *Proceedings of the 47th Annual Meeting of the Chicago Linguistic Society*. Chicago, IL: University of Chicago